Scientific Research and Essays Vol. 5(13), pp. 1583-1588, 4 July, 2010
Available online at http://www.academicjournals.org/SRE
ISSN 1992-2248 ©2010 Academic Journals

Full Length Research Paper

Drilling rate prediction using an innovative soft computing approach

Hamidreza Moradi*, Mohammad Hasan Bahari, Mohammad Bagher Naghibi Sistani and Aboozar Bahari

Ferdowsi University of Mashhad, Iran.

Accepted 17 February, 2010

An accurate approach for drilling rate prediction using a new soft computing approach is introduced in this paper. Drilling rate prediction is an important issue due to its crucial role in minimizing drilling cost. However, a large number of unforeseen factors and events influence the drilling rate and make it a complex and stochastic process and consequently difficult to predict. Many different techniques have been introduced for this task. Among those, Bourgoyne and Young model and its extensions have been widely used in drilling rate prediction during last decades. However, they did not provide satisfactory accuracy. In this research, a new soft computing approach is proposed over this problem and predicts the drilling rate with acceptable accuracy. Our practical data sets are nine wells of an Iranian gas field called "Khangiran". Simulation results show that the proposed intelligent approach is superior to the conventional methods in drilling rate prediction accuracy.

Key words: Rate of penetration, simulated annealing Fuzzylogic, drilling rate prediction, K-mean clustering, soft computing.

INTRODUCTION

The rate of penetration (*ROP*) or drilling rate is the speed at which a drill bit breaks the rock under it to deepen the borehole. Since drilling rate prediction is essential for drilling parameters determination and drilling cost optimazation, it has been a great concern for drilling engineers during last decades (Kaiser, 2007; Bourgoyne et al., 2003).

Rate of penetration is affected by many parameters such as hydraulics, weight on bit, rotary speed, bit type, mud properties etc (Akgun, 2007). Due to the uncountable uncertain factors influencing the drilling rate, unfortunately, there exists no exact mathematical relation between drilling rate and its parameters. Furthermore, their relationship to each other and to drilling rate is nonlinear and complex (Ricardo et al., 2007). However, experts have suggested some simplified models for mapping important variables of drilling on drilling rate (Kaiser, 2007). One of them is Bourgoyne and Young Model (BYM), which is used widely in practice (Bourgoyne and Young, 1974).

Bourgoyne and Young succeed to model the effect of different drilling parameters involving drilling rate as eight mathematical functions. For instance, the first function represents the effect of formation strength, bit type, mud type and solid content. In this model, there are some unknown parameters or coefficients which must be determined based on previous drilling experiences in the field. Bourgoyne and Young, (1974) suggested multiple regression method for this task. However, this method does not guarantee reaching physically meaningful coefficients (Bahari et al., 2008). To clarify, computed coefficients using multiple regression method can be negative or zero. Negative or zero coefficients are not physically meaningful. For instance, if the weight on bit coefficient be negative, it illustrates that increasing the weight on bit decreases the penetration rate; or if this value be zero, it means that increasing the weight on bit has no effect on

*Corresponding author. E-mail: Hamidreza.moradi@ieee.org. Tel: +98-511-8930606

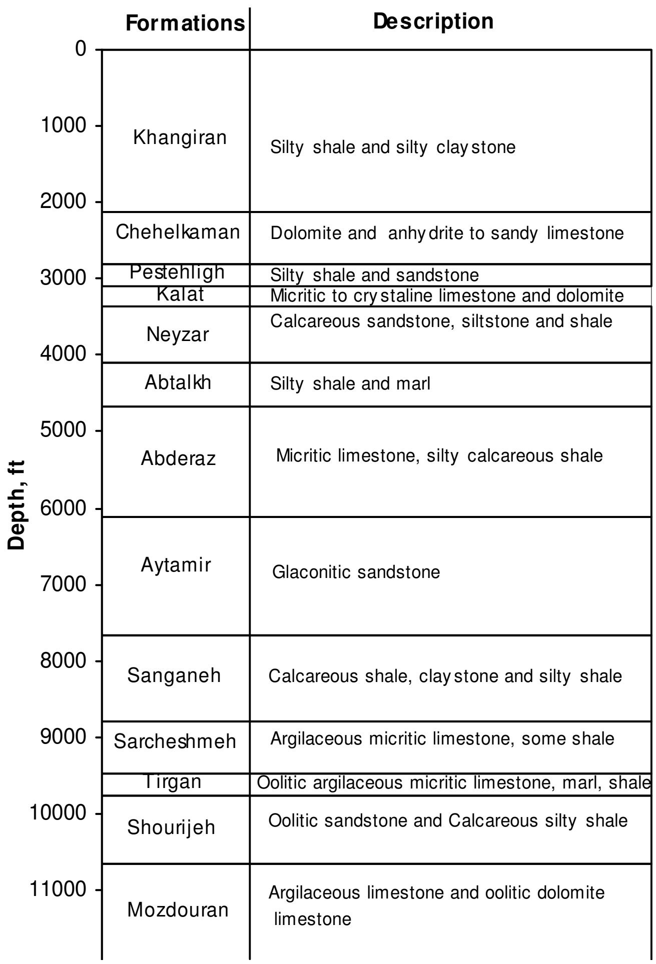

Figure 1. Stratigraphy column of a typical well in Khangiran field and formations description.

the drilling rate.

To reach meaningful results, some other methods such as non-linear least square data fitting with trust-region (Bahari and Baradaran, 2007) have recently been applied. Although extensions of BYM yield physically meaningful coefficient, they do not represent desired prediction accuracy.

On the other hand, soft computing methods have been successfully applied to different applications in the field of petroleum industry such as reservoir characterization (Zellou and Ouenes, 2007), optimum bit selection (Yilmaz et al., 2002), trap quality evaluation (Shi et al., 2004), and drilling rate prediction (Bahari and Baradaran, 2009) during past decades. In fact, these intelligent methodologies have many features that make them attractive to use in these problems. Among these, however, the ability to deal with ill-defined and noisy real signals and datasets are the most important one.

As mentioned before, drilling rate is affected by many uncertain parameters. To reach required accuracy, a novel drilling rate predictor is presented in this paper based on a Fuzzy system associated with Simulated Annealing (SA) in which receives the main drilling variables as inputs and predicts the rate of penetration (*ROP*) as output.

The rest of the paper is organized in the following manner: first we introduce the Khangiran Iranian gas field; then, the intelligent drilling prediction method is proposed; the simulation results are presented in section 4 and; the paper ends with conclusions in section 5.

KHANGIRAN IRANIAN GAS FIELD

Khangiran gas field is located in the northeast of Iran. This field was surveyed in 1937. In 1956, the stratigraphy plan was prepared and it was named in 1962. Figure 1 indicates the stratigraphy column and geological description of each formation for a typical well in this field. Khangiran field includes three gas reservoirs:

1. Mozdouran: The existence of sour gas in this reservoir was proved in 1968 and the production was started in 1983. It consists of thick layer limestone. Up till now, 37 wells have been drilled.
2. Shourijeh B: This reservoir was explored in 1968 and production was started in 1974. Shourijeh formation is mainly formed from sandstone layers. So far, seven wells have been drilled and completed in the reservoir. The gas from this reservoir is sweet and H2S free.
3. Shourijeh D: This reservoir was explored in 1987 and after drilling the well, production was started in the same year. Seven wells have been drilled up to now. The gas from this reservoir is sweet too.

FUZZY HYBRID SIMULATED ANNEALING *ROP* PREDICTOR

Wang has proved that certain classes of Fuzzy systems have universal approximation capability (Wang, 1997). He proved that for any given real continuous function $g(x)$ and any arbitrary $\varepsilon > 0$, there exists a Fuzzy system $f(x)$ such that:

$$\sup_{x \in U} \left| f(x) - g(x) \right| < \varepsilon \quad (1)$$

We consider *ROP* as a nonlinear function with eight inputs, $g(x)$. The inputs of g(x) are true vertical depth (D), weight on bit (W), bit diameter (d_b), rotary speed (N), pore pressure gradient (g_p), equivalent mud density (ρ_c), fractional bit tooth wear (h), jet impact force (F_j). Note that influential parameters and elements on the drilling rate are not limited to afore-mentioned eight parameters. However, these eight parameters are the most important and effective ones. First, we partition a typical Fuzzy system with undetermined parameters. Then, SA is utilized to determine different parameters of Fuzzy system and reach to needed estimator $f(x)$. At the other words, Fuzzy hybrid SA algorithm provides an estimator $f(x)$ to approximate $g(x)$ with minimum error.

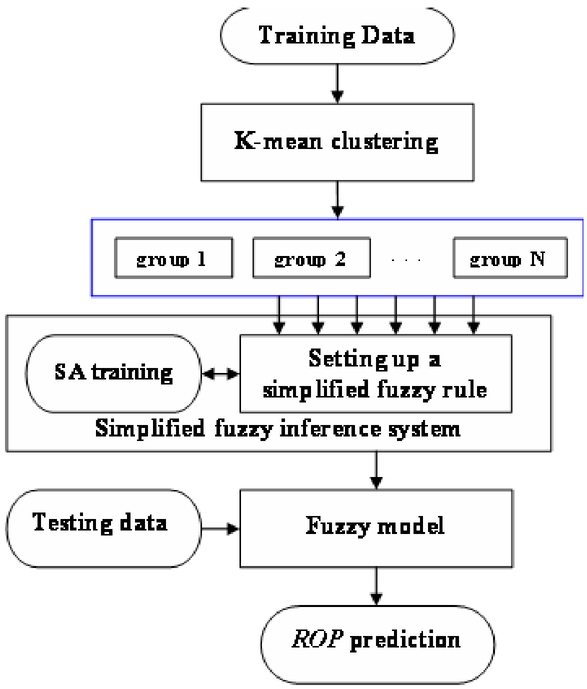

Figure 2. Framework of prediction procedure.

The architecture of the hybrid predictor system is discussed in this section. The overall framework of prediction procedure is illustrated in Figure 2. It can be interpreted from this figure that the predictor system is configured in three main steps as follows and are subsequently elaborated upon:

1. Clustering the training date in N groups (in this case 8 group) using k-mean clustering approach.
2. Setting up a typical Fuzzy system.
3. Using SA in order to determine the different parameters of designed Fuzzy system for achieving required estimator $f(x)$.

Clustering the training data

Choosing an appropriate number of rules is important in designing any Fuzzy system. To clarify, too many rules result in a complex Fuzzy system that may be unneces-sary for the problem, whereas too few rules produce a less powerful Fuzzy system that may be insufficient to achieve the objective.

We view the number of rules in the Fuzzy system as a design parameter and determine it based on the input-output pairs. The basic idea is to group the input-output pairs into clusters and to use one rule for one cluster; that is, the number of rules equals the number of clusters. The needed clusters are provided by K-mean clustering approach, which is a non-hierarchical clustering technique (Chang and Liu, 2008).

Fuzzy system

The main strategy of this research is to predict the ROP using a Fuzzy system. No definite membership function is defined antecedently, each cluster defines a rule and each data point defines a membership function. These membership functions are placed optimally using SA. A common rule of the Fuzzy system is represented as follows:

If x_1 **is** A_1^l **and** x_2 **is** A_2^l **and … and** x_8 **is** A_8^l **Then** y **is** B^l

Where, A_i^l and B^l are Gaussian with the following membership grade $h_i^l(x_i)$.

$$h_i^l(x_i) = \exp\left[-\frac{1}{2}\left(\frac{x_i - c_i^j}{\sigma_i^l}\right)^2\right] \qquad (2)$$

Where, c_i^l and $\sigma_{i_i}^{l}$ are mean and the standard deviation of Gaussian membership function for i^{th} input variable of l^{th} Fuzzy rule, respectively.

SA is used to determine c_i^l and σ_i of all membership functions of the Fuzzy system to reach the desired estimator $f(x)$. It is important to note that singleton fuzzifier and center average defuzzifier are used in this Fuzzy system.

Determining the parameters of Fuzzy system using SA

The main principle of applying the SA is finding a set of optimized parameters for the Fuzzy system and obtaining the required estimator $f(x)$. Figure 3 facilitates to show the operation process of SA. The process is briefly explained as follows:

Step 1

Initializing the parameters of SA, including the initial temperature, cooling coefficient, searching time of each temperature and the termination condition.

Step 2

A set of current solutions X is randomly gene-rated, which contains the mean and standard deviation of Gaussian membership function for 8 above-mentioned variables with no boundaries; the predicted MSE value of X can be obtained by the simplified Fuzzy inference system.

Step 3

Randomly search for a neighbor solution set X', which equals to X augments. All the neighbor solutions are substituted into the simplified Fuzzy inference system to have the predicted MSE value of X', and then the target distance (energy distance) of the current and neighbor solutions will be calculated as follows:

$$\Delta E = MSE(X') - MSE(X) \qquad (3)$$

If $\Delta E \leq 0$, then the current solution set will be replaced by the neighbor solution set; otherwise (when $\Delta E > 0$), the winning probability of the neighbor solution set is:

$$F(X') = \exp(-\Delta E / T) \qquad (4)$$

A random number u $\in$ (0,1) will be afterwards generated and if u < $F(X')$ the neighbor solution set will replace the current solution set;

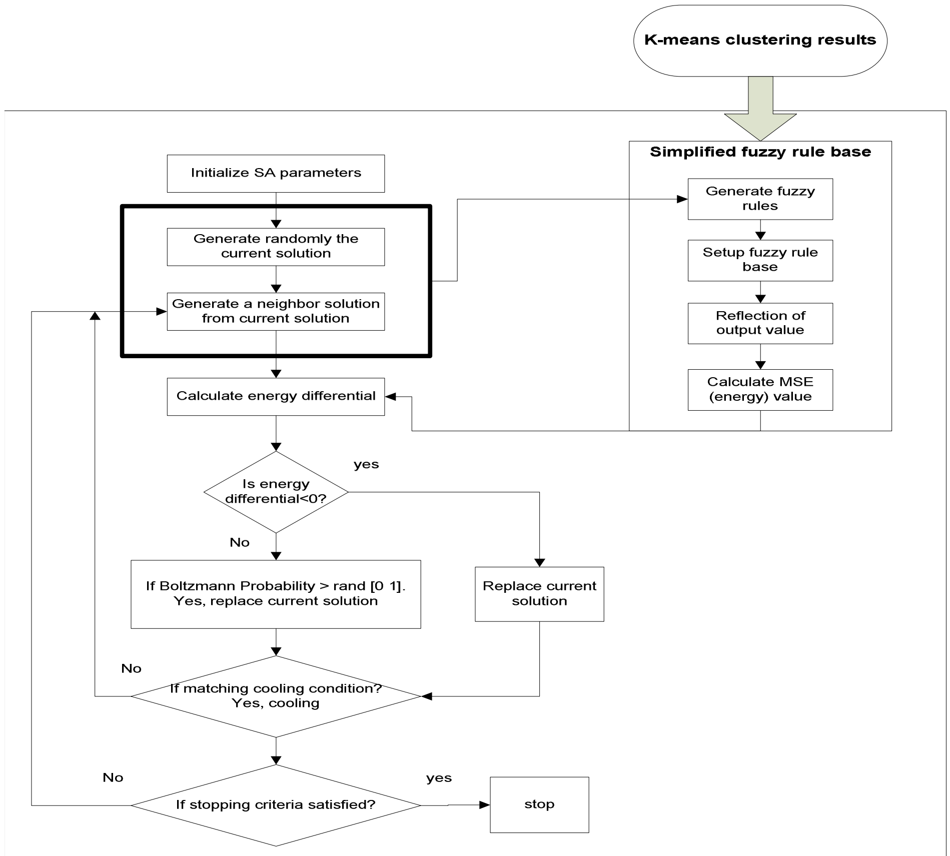

Figure 3. The operation process of simulated annealing approach.

otherwise, go back to step 3 to relocate the neighbor solution set, and add the searching time by 1.

Step 4

Compare X with the optimal solution set X''. If X is better, then it can be used to replace X''.

Step 5

If the maximum searching time of a certain temperature is achieved, then the temperature has to be cooled down, and the optimal solution of the last temperature will be set as the initial solution of the new temperature. If the maximum searching time is not achieved; go back to step 2.

Step 6

Check whether the termination condition is reached. If yes, finish the algorithm; otherwise, go back to carry out step 3 until the termination condition is fulfilled.

SIMULATION RESULTS

In this section, the proposed intelligent approach is compared with a conventional method namely GA aided BYM (Bahari and Baradaran, 2007). The simulation includes two phases. Training phase and testing phase.

Training phase

In the training phase, the following procedure is performed:

1. The daily drilling progress reporting different drilling parameters of 10 drilled wells (from the surface to the final reservoir depth) in this field were gathered initially. After data quality control, nine wells having more accurate data were adopted.
2. A database was constructed from available data of nine wells. The database includes quantities of D, W, d_b, N, g_p, ρ_c, h, F_j and achieved ROP in each formation. It must be noted that the fractional tooth wear is expressed just at the end of bit running. Therefore, only drilling data at ending the bit run can be used. Table 1 provides a sample of the required data, which is included in our database.
3. For determining the different parameters of Fuzzy system using SA, a training data set is formed by randomly choosing 75% of available data for each formation. The

Table 1. A sample of required data obtained from wells daily drilling progress reports.

Well No.	ROP (ft/hr)	D (ft)	W (1000lbf)	d_b (in.)	N (Rpm)	ρ_c (lbm/gal)	h (%)	g_p (lbm/gal)	F_j (lbf)
Well 1	41.5	1411	15	17.5	130	9.96	0.25	8.62	1776
Well 2	24.3	359	15	26	130	8.95	0.25	7.62	1611
Well 3	7.3	1772	7.5	17.5	110	10.3	0.25	8.95	1185
Well 4	9.5	1969	10	17.5	110	10.8	0.5	9.49	1324
Well 5	5.7	1900	9	17.5	100	10.5	0.5	9.15	1186
Well 6	25.9	1575	15	17.5	90	10.4	0.38	9.09	2196

Table 2. The estimation accuracy of the proposed method in comparison with GA aided BYM.

Formation	Proposed scheme MSE of estimation	GA aided BYM MSE of estimation	Improvement (%)
Sanganeh	2.1	4.2	50
Chehelkaman	2.35	4.22	44.3
Sarcheshmeh	0.99	2.35	57.8
Shourijeh	.88	1.8	51.1
Mozdouran	1.0	1.3	23
Neyzar	2.0	4.4	54.4
Aytamir	2.73	5.2	47.5
Abderaz	3.63	6.55	44.5
Total	15.68	30.02	47.77

remained data is used to test the proposed method.
4. The training data set for each formation is clustered by applying k-mean clustering approach into eight clusters.
5. A typical Fuzzy system with mentioned structure and undetermined parameters is partitioned for each formation.
6. Using the selected training data sets of each formation, the SA is employed to determine the parameters of partitioned Fuzzy systems. Note that *ROP* is dependant to local drilling conditions and must be computed for each formation using prior drilling data obtained in the area. Consequently, Fuzzy systems of each formation must be trained independently.

Using the above-mentioned procedure, a hybrid predictor is obtained for each formation of the field.

Testing phase

In order to test the proposed intelligent predictor, the following procedure is performed:

1. A testing data set is formed using 25% of all available data for each formation. Note that the testing data set was not used in training phase.
2. By applying values of D, W, N, d_b, g_p, ρ_c, h, and F_j in each formation to the hybrid predictor, values of *ROP* are calculated.
3. Mean squared error (MSE) of *ROP* estimation is calculated for each formation.

We repeated the training and testing phases for 1000 times. The mean of computed values of the third step in testing phase over 1000 times are indicated in Table 2. For comparison purpose, GA aided BYM is implemented on the same data set and results obtained from GA aided BYM is demonstrated too. It can be interpreted that the proposed scheme is more accurate than conventional one as it presents more accuracy.

Conclusion

For drilling cost optimization, one important issue, which is aimed at is accurate drilling rate prediction. A commonly used method for drilling rated prediction is Bourgoyne and Young model (BYM) and its different extensions. BYM provides a mapping of the main drilling variables on the drilling rate. However, it does not represent enough accuracy. The main contribution of this paper is to introduce a new method based on Fuzzy logic to predict the drilling rate accurately. Simulation results over nine wells of Khangiran Iranian gas field confirm the effectiveness of new intelligent method.

ACKNOWLEDGMENT

The authors would like to thank the drilling staff of Iranian

Central Oil Fields Company (I. C. O. F. C) for their contributions and cooperation.

REFERENCES

Akgun F (2007). Drilling rate at the technical limit. Int. J. Petroleum Sci. Tech., 1: 99-118.

Bahari A, Baradaran SA (2007). Trust-region approach to find constants of Bourgoyne and Young penetration rate model in Khangiran Iranian gas field. Proceeding of SPE Latin American and Caribbean Petroleum Engineering Conf., Buenos Aires, Argentina, pp. 15-18.

Bahari A, Baradaran Seyed A (2009). Drilling cost optimization in a hydrocarbon field by combination of comparative and mathematical methods. Pet. Sci., 6: 451-463.

Bahari MH, Bahari A, Nejati Moharrami F, Naghibi-S MB (2008). Determining Bourgoyne and Young Model coefficients Using Genetic Algorithm to Predict Drilling Rate. J. Appl. Sci., 8(17): 3050-3054.

Bourgoyne AT, Millheim KK, Chenevert ME, Young FS (2003). Applied Drilling Engineering. 9^{th} ed., SPE, Richardson, TX, p. 232.

Bourgoyne AT, Young FS (1974). A multiple regression approach to optimal drilling and abnormal pressure detection. Trans. SPEJ, p. 371.

Chang PC, Liu CH (2008). A TSK type Fuzzy rule base system for stock prediction. Expert Systems with applications. 34:135-144.

Kaiser MJ (2007). A survey of drilling cost and complexity estimation models. Int. J. Petroleum Sci. Tech., 1: 1-22.

Ricardo J, Mendes P, Fonseca TC, Serapiao ABS (2007). Applying a genetic neuro-model reference adaptive controller in drilling optimization. World Oil Magazine, p. 228.

Wang LX (1997). A Course in Fuzzy Systems and Control, Prentice-Hall, NJ, pp. 118-140.

Yilmaz Y, Demircioglu C and Akin S (2002). Application of Artificial Neural Networks to optimum bit selection. Comp. Geosci., 28(2): 261-269.

Zellou AM and Ouenes A (2007). Integrated fractured reservoir characterization using neural networks and Fuzzy logic: three case studies. J. Petroleum geol., 24(4): 459-476.